

### Mehrwertdienste für das Information Retrieval: das Projekt IRM

Mayr, Philipp; Mutschke, Peter; Schaer, Philipp; Sure, York

Postprint / Postprint

Konferenzbeitrag / conference paper

**Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:**

GESIS - Leibniz-Institut für Sozialwissenschaften

#### Empfohlene Zitierung / Suggested Citation:

Mayr, P., Mutschke, P., Schaer, P., & Sure, Y. (2013). Mehrwertdienste für das Information Retrieval: das Projekt IRM. In H. P. Ohly (Hrsg.), *Wissen - Wissenschaft - Organisation: Proceedings der 12. Tagung der Deutschen Sektion der Internationalen Gesellschaft für Wissensorganisation* (S. 131-139). Würzburg: Ergon Verl. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-338870>

#### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC Lizenz (Namensnennung-Nicht-kommerziell) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by-nc/4.0/deed.de>

#### Terms of use:

This document is made available under a CC BY-NC Licence (Attribution-NonCommercial). For more information see: <https://creativecommons.org/licenses/by-nc/4.0>

# Mehrwertdienste für das Information Retrieval

## Das Projekt IRM

Philipp Mayr, Peter Mutschke, Philipp Schaer, York Sure  
GESIS - Leibniz-Institut für Sozialwissenschaften, Bonn

### Zusammenfassung

Ziel des Projekts ist die Entwicklung und Erprobung von metadatenbasierten Mehrwertdiensten für Retrievalumgebungen mit mehreren Datenbanken: a) Search Term Recommender (STR) als Dienst zum automatischen Vorschlagen von Suchbegriffen aus kontrollierten Vokabularen, b) Bradfordizing als Dienst zum Re-Ranking von Ergebnismengen nach Kernzeitschriften und c) Autorenzentralität als Dienst zum Re-Ranking von Ergebnismengen nach Zentralität der Autoren in Autorenetzwerken. Schwerpunkt des Projektes ist die prototypische Implementierung der drei Mehrwertdienste in einer integrierten Retrieval-Testumgebung und insbesondere deren quantitative und qualitative Evaluation hinsichtlich Verbesserung der Retrievalqualität bei Einsatz der Mehrwertdienste.

### 1. Einleitung

Wissenschaftsportale wie vascoda (<http://www.vascoda.de>) und sowiport ([www.sowiport.de](http://www.sowiport.de)) fassen die Informationsangebote mehrerer Anbieter zu einer einheitlichen Recherche zusammen. Dahinter steht der Anspruch, für die wissenschaftliche Recherche auf der Basis qualitativ erschlossener und kontrollierter Informationen deutlich bessere Ergebnisse zu liefern als die generellen Web-Suchmaschinen oder auch Google Scholar. Erfahrungen im vascoda-Kontext verweisen jedoch auf zwei wesentliche Schwachstellen fachübergreifender förderierter Portale (vgl. Mayr et al., 2008):

- Benutzer wählen in vielen Fällen ihre Suchterme frei aus dem Gedächtnis. Diese können somit gar nicht oder nur zufällig einem kontrollierten Vokabular zugeordnet werden, so dass auch Heterogenitätskomponenten, die automatisch zwischen verschiedenen Fachvokabularen umschlüsseln, nicht stringent funktionieren.
- Die in vascoda und sowiport eingesetzten, im Retrieval geläufigen Ranking-Faktoren auf der Basis von Termhäufigkeiten führen im speziellen Kontext der mit kontrolliertem Vokabular erschlossenen Dokumente nicht automatisch zu befriedigenden Ergebnissen.

Ziele des Projekts sind daher die Erprobung und Weiterentwicklung von Term-Vorschlagssystemen auf der Basis kontrollierter Vokabulare als Lösungsstrategie für das Problem der Wahl der „richtigen“ Suchterme. Der im Projekt eingesetzte Search Term Recommender (STR) beruht auf der statistischen Assoziation von Suchbegriffen mit Termen und Phrasen aus dem kontrollierten Vokabular (s. Petras, 2006). Hierzu wird eine Kookkurrenzanalyse zwischen „freien“ Begriffen aus Titel oder Abstract und Deskriptoren aus dem kontrollierten Vokabular durchgeführt. Je öfter ein freier Begriff mit einem Deskriptor zusammen erscheint, umso wahrscheinlicher ist die Annahme, dass diese zwei Terme semantisch miteinander assoziiert sind.

Als Lösungsstrategie für das Rankingproblem werden zwei aus der Scientometrie und Netzwerkanalyse entlehnte Verfahren getestet, die eine ganz andere, nicht-textorientierte Art des Re-Ranking vorsehen.

- **Bradfordizing:** Das Bradfordizing (White, 1981) sortiert bzw. rankt nach Kernzeitschriften. Die Zeitschriften zu einer Suchanfrage werden gemäß ihrer Häufigkeit (Anzahl der Zeitschriftenaufsätze) im Suchergebnis gelistet. Aufsätze von sogenannten Kernzeitschriften werden beim Bradfordizing vor die Zeitschriftenartikel geordnet, die nur durchschnittlich oder selten Artikel zu dem gesuchten Thema enthalten. Das Bradfordizing kann insofern als ein innovatives Re-Ranking Verfahren verstanden werden, als es für eine beliebige Suchanfrage über eine nachgeordnete Umorganisation des Suchergebnisses zuverlässig (hohe Robustheit dieser Häufigkeitsverteilung) die Publikationen aus Kernzeitschriften für eine übergreifende Suche an den Anfang der Ergebnisliste befördert.
- **Autorennetzwerke:** Der Ansatz des Autorennetzwerkmodells (Mutschke, 2004) ist es, Wissen über das Interaktions- und Kooperationsgeschehen in einer Scientific Community und den (sozialen) Status ihrer Akteure (Zentralität) für die Informationssuche auszunutzen. Das Modell identifiziert zentrale Autoren innerhalb einer Dokumentenkollektion bzw. Treffermenge einer Recherche und rankt diese nach Maßgabe der Zentralität der Autoren im Autorennetzwerk, so dass dem Benutzer Publikationen zentraler Autoren zuerst angezeigt werden.

Bei allen drei Diensten stellt sich jedoch die Frage, ob sie Precision und Recall der Ergebnismenge erhöhen und ob sie von den Nutzern eines Portals verstanden werden bzw. ob sich der Mehrwert der Verfahren intuitiv erschließt. Ein Schwerpunkt des Projekts ist daher die Evaluation der Retrievalqualität und Benutzerakzeptanz der drei Mehrwertdienste bei einer Suche in einem datenbankübergreifenden Rechercheportal. Von besonderem Interesse ist dabei, ob und inwieweit sich die Retrievalqualität bei einer Kombination der Mehrwertdienste verbessert.

## 2. Datenbanken im Projekt

Als Dokumentbasis für das Prototypensystem und die Evaluation der Mehrwertdienste wurden Datenbanken ausgewählt, die hinsichtlich ihrer Erschließungstiefe und Metadatenstruktur die methodischen Anforderungen der zu evaluierenden Mehrwertdienste erfüllen. Nach entsprechender Analyse der potentiell zur Verfügung stehenden Datenbanken wurden als besonders geeignet für die Evaluation und Erprobung der drei Mehrwertdienst ausgewählt:

- SOLIS (GESIS): deutschsprachige sozialwissenschaftliche Literatur, 384.000 Nachweise
- SPOLIT (Bundesinstitut für Sportwissenschaft): sportwissenschaftliche Literatur des In- und Auslandes, 162.737 Nachweise
- FIS Bildung (Deutsches Institut für Internationale Pädagogische Forschung): erziehungswissenschaftliche Literatur, 697.528 Nachweise
- SA (Cambridge Scientific Abstracts): Sociological Abstracts, 376.191 Nachweise
- PEI (Cambridge Scientific Abstracts): Physical Education Index, 127.754 Nachweise
- FES (Friedrich-Ebert-Stiftung): Bibliothekskatalog, 487.559 Nachweise.

Für die Auswahl dieser Datenbanken sprechen folgende Gründe:

- SOLIS, FIS Bildung und SPOLIT haben ein mittelgroßes kontrolliertes Vokabular mit einer sehr hohen Indexierungstiefe (10 Deskriptoren pro Dokument). SA und PEI haben eher kleinere Vokabulare mit einer mittleren Indexierungstiefe (ca. 5 Deskriptoren pro Dokument). FES hat demgegenüber eine niedrige Indexierungstiefe (unter 2 Deskriptoren pro Dokument). Damit sind Datenbanken unterschiedlicher Indexierungstiefe und -breite vertreten sowie mit Deutsch (SOLIS, FIS Bildung, SPOLIT, FES) und Englisch (SA, PEI) auch zwei Sprachen.

- In den ausgewählten Datenbanken sind i.d.R. ISSN- oder ISBN-Nummern repräsentiert. Datensätze ohne diese Identifier können für das Verfahren Bradfordizing nur in Ausnahmen berücksichtigt werden.
- In allen Datenbanken sind i.d.R. volle Ansetzungen der Autorennamen sowie Koautorenschaften repräsentiert, sodass sich diese Datenbanken grundsätzlich für eine Autorennetzwerkanalyse eignen.

Für die Auswahl von SOLIS/SA und SPOLIT/PEI sowie FIS Bildung spricht außerdem, dass diese Datenbanken unterschiedliche fachwissenschaftliche Disziplinen, nämlich Sozialwissenschaften (SOLIS, SA), Sportwissenschaften (SPOLIT, PEI) und Pädagogik (FIS Bildung) repräsentieren, sodass auch Unterschiede zwischen den Fächern evaluiert werden können. Insgesamt sind rund 2.19 Millionen Dokumente über den Prototypen indexiert.

### 3. Anwendung der Mehrwertdienste

Das Grundszenario für eine Integration der drei Mehrwertdienste ist die Manipulation der Ergebnismengen durch kombinierte Anwendung der Mehrwertdienste. Hier sind zwei Grundvarianten zu unterscheiden:

- die iterative Anwendung mehrerer Mehrwertdienste auf die Ergebnismenge, wobei ein Mehrwertdienst als Filter für den nächsten fungiert: Eine z. B. per STR erzeugte Ergebnismenge wird per Bradfordizing so umsortiert, dass Beiträge in Kernzeitschriften hochgerankt werden. Auf die Menge der Beiträge in Kernzeitschriften wird wiederum Autorenzentralität angewendet, so dass schließlich Beiträge zentraler Autoren in Kernzeitschriften an der Spitze der Ergebnisliste stehen.
- die simultane Anwendung mehrerer Mehrwertdienste auf die Ergebnismenge und Schnittmengenbildung: Auf eine z. B. per STR erzeugte Ergebnismenge wird unabhängig voneinander Bradfordizing und Autorenzentralität angewendet und anschließend die Schnittmenge der Ergebnismengen gebildet; es stehen ebenfalls Beiträge zentraler Autoren in Kernzeitschriften an der Spitze, die Einzel-Rankings erfolgen allerdings auf Basis jeweils anderer Ausschnitte aus der Ergebnismenge als beim iterativen Verfahren. Eine Schwierigkeit dieser Variante ist allerdings die „richtige“ Integration der Einzel-Rankings zu einem Gesamt-Ranking

Für beide Varianten sind unterschiedliche Zweier- und Dreier-Kombinationen möglich. Insgesamt ergeben sich folgende Einsatzmöglichkeiten (s. Tabelle 1):

Kombination	Autorennetzwerke	Bradfordizing	STR
0	-	-	-
1	1	-	-
2	-	1	-
3	-	-	1
4	1	2	-
5	1	1	-
6	2	1	-
7	2	-	1
8	-	2	1
9	2	3	1
10	3	2	1
11	2	2	1

Tabelle 1: Mögliche Einsatzkombinationen. Zahl bedeutet Reihenfolge des Einsatzes. (Legende: 1. Zeile (0) = Baseline, - = kein Einsatz). Die gleiche Zahl bedeutet simultaner Einsatz mit Ergebnis-Integration.

Im Projekt werden für die Evaluation die Einzelverfahren sowie Zweier-Kombinationen derselben priorisiert (z.B. Mehrwerteffekte durch Kombination von Bradfordizing und Autorenzentralität). Dabei wird davon ausgegangen, dass beim Re-Ranking die Ermittlung der Rankingwerte auf der Basis der jeweiligen Ergebnismengen erfolgen (Ex-Post-Ranking-Ansatz). Eine Alternative wäre ein Ex-Ante-Ranking, bei dem die Rankingwerte auf Basis der Gesamtkollektion berechnet und für die Indexierung der Dokumente verwendet werden. Die Berücksichtigung von Ex-Ante-Varianten würde die Zahl der Kombinationsmöglichkeiten allerdings entsprechend erhöhen. Im Projekt wird die Ex-Post-Variante favorisiert, da es auf thematischen Suchergebnissen basiert.

### Suchexpandierung (STR)

Bei der Entwicklung von Anzeige- und Browsing-Strukturen für die STR-Termvorschläge spielt die Frage der Nutzerinteraktion eine zentrale Rolle. Hier ist zu entscheiden, inwieweit der Benutzer in den Termvorschlagsprozess eingebunden wird oder ob die Auswahl vollautomatisch im Hintergrund abläuft – ohne dass der Benutzer davon erfahren würde. Drei Varianten sind möglich:

1. Die Begriffe werden automatisch ausgewählt, keine Benachrichtigung des Benutzers,
2. die Begriffe werden automatisch ausgewählt, Benachrichtigung des Benutzers oder
3. der Benutzer wählt Begriffe aus.

Die erste Möglichkeit der Einbindung ist es, den STR vollkommen autonom im Hintergrund agieren zu lassen, ohne den Benutzer entsprechend darüber in Kenntnis zu setzen. In diesem Fall würden vom STR-Modul die n ersten Begriffe zur Suchanfrage hinzugefügt. Es gibt die Möglichkeit, den Benutzer zusätzlich über die vom System

vorgenommenen Veränderungen zu unterrichten. Dieses Verfahren ist analog zur Funktionsweise der jetzigen Implementation des Heterogenitätsservices in sowiport. In der Diplomarbeit von Daniela Holl wurde die Variante drei umgesetzt (s. Holl, 2009), bei der einem Benutzer nach der Eingabe von freien Begriffen in der Suchmaske in einem weiteren Schritt Begriffe vorgeschlagen werden. Fraglich aus Sicht der Gebrauchstauglichkeit ist hier, ob Benutzer gewillt sind, einen weiteren Interaktionsschritt in Kauf zu nehmen. Generell lässt sich für die Frage der Nutzerinteraktion und Anzeige sagen, dass es sich hier um einen kontinuierlichen Raum zwischen beiden extremen Positionen der vollkommenen Autonomie der Software und der Kontrolle durch den Benutzer handelt. Die drei vorgestellten Varianten erlauben diverse weitere Kombinations- und Abstufungsmöglichkeiten. So wäre denkbar, dass in einem iterativen Suchprozess zunächst automatisch vom System Begriffe vorgeschlagen werden, die nachträglich noch vom Benutzer angepasst oder bewertet werden können.

### **ReRanking**

Für den Einsatz eines Re-Rankings mit Bradfordizing und/oder Autorenzentralität sind folgende Szenarios denkbar:

1. pro-aktives Re-Ranking der Ergebnismenge (ohne Benutzerinteraktion)
2. Re-Ranking der Ergebnismenge nach Aufforderung durch den Benutzer (Klick auf Button oder Auswahlliste),
3. Listing der n relevantesten Autoren bzw. Kernzeitschriften und Einschränkung der Ergebnismenge auf die Dokumente der vom Benutzer ausgewählten Items oder
4. Verlinkung der Kernzeitschriften mit zentralen Autoren.

Diese vier Szenarios sind auch in kombinierten Variationen denkbar. So können nach dem ersten Listen der Dokumente über Auswahlen (Selektion durch den Nutzer) weitere Untermengen des Ergebnisses gebildet werden, die dann mit weiteren Verfahren gekoppelt werden. Dabei wäre sowohl der Ansatz ex-ante (Vorberechnung der drei Zonen für die gesamte Datenbank oder mehrere Datenbanken) als auch ex-post (zur Laufzeit einer Suchanfrage) möglich. Im Rahmen eines Usability-Tests soll evaluiert werden, welche der genannten Szenarios in einer Retrievalumgebung prioritär umgesetzt werden sollten.

## **4. Implementierung der Mehrwertdienste**

Ziel ist es eine erste prototypische Umsetzung der einzelnen Verfahren zu realisieren um anschließend Usability- und Retrievaltest mit dem System durchführen zu können. Hierfür wurde auf das freie Web-Framework Grails zurückgegriffen, welches auf quelloffenen Standardlösungen wie Spring, Hibernate oder Sitemesh basiert. Grails ist vollständig in die Java-Plattform integriert, da als Programmiersprache auf Groovy gesetzt wird. Über einen umfangreichen Plugin-Mechanismus lassen sich dabei weitere Elemente und Funktionalitäten (z. B. AJAX-Unterstützung, Nutzerverwaltung etc.) nachrüsten.

Neben dem Grails Framework wird zur Suche und zur Datenhaltung auf den freien Enterprise-Suchserver Solr zurückgegriffen. Dieser basiert auf der Suchbibliothek Lucene und bietet u. a. die Möglichkeit zur direkten Facettierung und einer Einbindung in beliebige eigene Projekte mittels eines RESTful Webservices, der neben XML z. B. auch JSON spricht. Auch Solr verfügt über einen eigenen Plugin-Mechanismus und ist durch seine Quelloffenheit beliebig anpassbar. Für Rerankingverfahren der Autorenzentralität wurde über diese Schnittstelle beispielsweise bei der Indexierung ein Vorberechnungsschritt ausgeführt und das Ergebnis zusammen mit den Metadaten in den Index geschrieben.

## The IRM Show Case - Powered by Grails

Query:   ☐ Debug mode

Suggest search terms

Automatic query expansion ☒

Controlled vocabulary

Sort suggested terms by

Rerank the result list (Sorry, only one at a time for the moment...)

First rerank method

Second rerank method

**Total hits: 1515**

[next 10 documents >>](#)

[Central journals \(expand/collapse list\)](#)

Search term suggestions [\[expand/collapse\]](#)  
[tag cloud](#)

**Autopoiesis** [Begriffsbildung](#)

**Erkenntnistheorie**

[Exclusion](#) [Feldtheorie](#) [Funktionalismus](#)  
[Gedächtnis](#)

**Gesellschaftstheorie**

[Habermas, J. Handlungstheorie](#)  
[Individualität](#) [Individuum](#) [Interaktion](#)  
[Kommunikation](#)  
[Kommunikationsmedien](#)

**Komplexität**

[Konstruktivismus](#)

**Kontingenz** [Kritik](#)

1. Jäger, Wieland; Schimank, Uwe (2005): *Organisationsgesellschaft : Facetten und Perspektiven in Hagener Studententexte zur Soziologie* (3-531-14336-0) [toggle abstract](#)
2. Lange, Stefan; Schimank, Uwe (2001): *A political sociology for complex societies: Niklas Luhmann in Blackwell companions to sociology* [toggle abstract](#)
3. Greshoff, Rainer; Schimank, Uwe (2006): *Integrative Sozialtheorie? : Esser - Luhmann - Weber* (3-531-14354-9) [toggle abstract](#)
4. Greshoff, Rainer; Kneer, Georg; Schimank, Uwe (2003): *Die Transintentionalität des Sozialen : eine vergleichende Betrachtung klassischer und moderner Sozialtheorien* (3-531-14037-X) [toggle abstract](#)
5. Rosewitz, Bernd; Schimank, Uwe (1988): *Verselbständigung und politische Steuerbarkeit gesellschaftlicher Teilsysteme in Schriften aus dem Max-Planck-Institut für Gesellschaftsforschung Differenzierung und Verselbständigung zur Entwicklung gesellschaftlicher Teilsysteme* (3-593-34030-5) [toggle abstract](#)
6. Czada, Roland; Schimank, Uwe (2000): *Institutionendynamik und politische Institutionengestaltung: Die zwei Gesichter sozialer Ordnungsbildung in Schriften aus dem Max-Planck-Institut für Gesellschaftsforschung Bd. 39* [toggle abstract](#)

**Gesellschaftliche Komplexität und kollektive Handlungsfähigkeit**

Abbildung 2: IRM-Prototyp

Für die Umsetzung des STR Prototypen wurde eine vorhandene Installation des kommerziellen Klassifikationssystems Recommind Mindserver verwendet. Die für die Funktionalität des STR notwendigen Assoziationsberechnungen konnten dabei vom Mindserver sehr effizient umgesetzt werden. Die Software setzt auf einen hybriden Ansatz von PLSA (Probabilistisch Latent Semantische Analyse) und Support Vector Machines. Diese erlauben es, mit den aus sowiport extrahierten Datenbeständen einen STR-Service zu implementieren. Der Mindserver wurde per Webservice in die o. g. Grails-Plattform integriert, sodass erste Anzeige- und Browsing-Strukturen in Form von Listen und Word-Clouds implementiert werden konnten (vgl. Abb. 2). Nach Eingabe von beliebigen Begriffen werden semantisch nahe Begriffe vom System vorgeschlagen, die wiederum für die Anfrage-Erweiterung genutzt werden können.

Die jetzige Systemarchitektur mit den Komponenten Grails, SOLR, Mindserver und der externen Datenbank xHive ist auf der Abbildung 3 zu sehen. Die xHive-Datenbank, in der die bibliografischen Metadaten noch umfangreicher gespeichert sind, ist im Moment noch nicht angebunden – dies kann aber bei Bedarf geschehen.

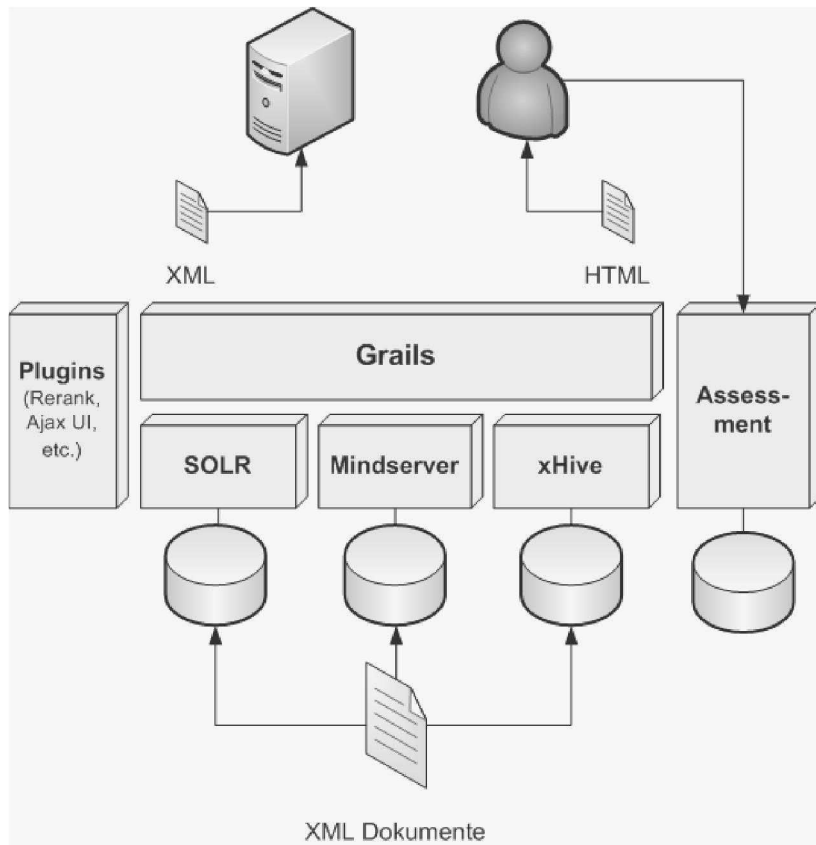


Abbildung 3: Architektur des IRM-Prototypen

## 5. Evaluation der Mehrwertdienste

Die Evaluation der Effektivität der implementierten Mehrwertdienste ist ein Schwerpunkt des Projekts. Im Zentrum der Evaluation stehen die folgenden Forschungsfragen:

- Verbessert sich das Retrieval durch Einsatz der Mehrwertdienste?
- Werden durch die Mehrwertdienste ANDERE relevante Treffer favorisiert als durch Standardverfahren? Wenn ja, wie hoch ist der (quantitative/qualitative) Mehrwert?
- Werden die Mehrwertdienste vom Benutzer verstanden und angenommen (Akzeptanz)?
- Gibt es eine (weitere) Verbesserung durch Kombination der Dienste? Wenn ja, welche Kombinationen sind am erfolgreichsten?

In diesem Zusammenhang werden mindestens zwei Kategorien der Effektivität unterschieden. Zum einen besteht an jeden Retrievaldienst der Anspruch, die Suche insofern zu verbessern, dass mehr relevante Dokumente für ein bestimmtes Informationsbedürfnis gefunden werden.

Der Einfluss der neu aufgesetzten Mehrwertdienste in diesem Sinne lässt sich traditionell über quantitative Retrieval-Evaluationsverfahren (Laboruntersuchungen) analysieren. Dazu werden Nutzeranfragen und deren Suchresultate mit und ohne Zunahme der zusätzlichen



Dienste verglichen und insbesondere die Precision (Anteil der relevanten Suchergebnisse aus allen Suchergebnissen) auf Basis von intellektuellen binären Relevanzbewertungen gemessen. Hier kann auf bereits evaluierte Fragestellungen aus den Projekten KoMoHe und CLEF zurückgegriffen werden (s. Mayr, 2009). Für diese Projekte liegen relevanzbewertete Dokumentsammlungen vor, so dass eine Evaluation der drei Mehrwertdienste auf Basis dieser Daten ohne Benutzerinteraktion möglich ist. Für das Projekt wurde ein Test der Precision unter den ersten 10, 20 und 30 Treffern der mit den zu untersuchenden Mehrwertdiensten generierten Ergebnismengen begonnen. Erste Ergebnisse zeigten einen erheblichen informationellen Mehrwert bei einem Re-Ranking mit Bradfordizing und Autorenzentralität. Darüber hinaus sollen Precision-Tests mit unterschiedlichen Benutzergruppen mit neuen Topics, die im Projekt zu definieren sind, und freien Recherchen durchgeführt werden.

Neben quantitativen Precision-Tests sollen bei der Evaluation der drei Mehrwertdienste auch qualitative Aspekte betrachtet werden, die den Einfluss der Mehrwertdienste auf das Sucherlebnis an sich (z. B. emotive Effekte), d. h. die Zufriedenheit der Nutzer mit dem Ablauf der Suche untersuchen. Dazu gehören Erfahrungen wie Schwierigkeiten in der Handhabung unterschiedlicher Retrievalinstrumente und -oberflächen (Usability-Tests) sowie die Gewissheit des Nutzers, mehrheitlich relevante Dokumente gefunden zu haben. Die erhöhte Schwierigkeit der Nutzerführung, Transparenz des Suchprozesses und der Suchergebnisse sowie die Visualisierung der Terme bzw. neugeordneten Dokumente soll analysiert werden. Weiterhin müssen auch veränderte Einsatzszenarien für den integrierten Einsatz der Mehrwertdienste analysiert werden. Es muss geprüft werden, ob die Kombination der Mehrwertdienste weitere positive Effekte generieren kann oder, im Gegenteil, zusätzliche Vorkehrungen getroffen werden müssen, um potentielle negative Effekte auszuschalten. Die Usability- und Benutzertests mit dem IRM-Prototypen sollen Anteile von aufgabengestützten Recherchen und freiem Explorieren der Mehrwertdienste enthalten.

Die Evaluation der Mehrwertdienste soll mit unterschiedlichen Benutzergruppen durchgeführt werden:

- Interessierte Öffentlichkeit (z. B. sowiport-Nutzer, Journalisten)
- Informationsexperten (Dokumentare, Bibliothekare)
- Fachwissenschaftler (Sozialwissenschaftler, Sportwissenschaftler, Bildungsforscher)
- Studenten-Gruppen (Hochschule Darmstadt, Uni Koblenz, HU Berlin)

Es ist geplant, Gruppen in der Größe von ca. 20 - 30 Personen zu untersuchen. Kandidaten für alle Gruppen und Bereiche wurden identifiziert. Die Evaluation kann direkt nach Abschluss der Implementierung der Mehrwertdienste beginnen.

Im Projekt sollen neben klassischen Dokumentlisten auch andere Formen von Ergebnismengen evaluiert werden (z. B. Varianten von Trefferlisten, Journal-Listen, Autorenlisten, Term-Listen). Zu diesem Zweck wurde ein Assessment-Tool entwickelt, das in der Lage ist, Relevanzurteile aus einer beliebigen Trefferliste zu speichern (siehe dazu Architektur-Abbildung 3).

## Literatur

- Holl, Daniela (2009): Search Term Recommender auf Basis der Recommind Software. Computervisualistik Universität Koblenz-Landau Koblenz. Diplomarbeit
- Mayr, Philipp (2009): Re-Ranking auf Basis von Bradfordizing für die verteilte Suche in Digitalen Bibliotheken. Philosophische Fakultät I, Institut für Bibliotheks- und Informationswissenschaft Humboldt-Universität zu Berlin. 237 p., URL: <http://edoc.hu-berlin.de/dissertationen/mayr-philipp-2009-02-18/PDF/mayr.pdf>
- Mayr, Philipp; Mutschke, Peter; Petras, Vivien (2008): Reducing semantic complexity in distributed digital libraries: Treatment of term vagueness and document re-ranking. In: Library Review 57, No. 3, pp. 213-224. URL: [http://www.ib.hu-berlin.de/~mayr/arbeiten/mayr-etal\\_LR08.pdf](http://www.ib.hu-berlin.de/~mayr/arbeiten/mayr-etal_LR08.pdf)
- Mutschke, Peter (2004): Autorennetzwerke: Verfahren der Netzwerkanalyse als Mehrwertdienste für Informationssysteme. Bonn: Informationszentrum Sozialwissenschaften. 48 p. (IZ-Arbeitsbericht Nr. 32) URL: [http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis\\_reihen/iz\\_arbeitsberichte/ab\\_32.pdf](http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/iz_arbeitsberichte/ab_32.pdf)
- Petras, Vivien (2006): Translating Dialects in Search: Mapping between Specialized Languages of Discourse and Documentary Languages. University of California, Berkeley Berkeley, USA. URL: <http://www.sims.berkeley.edu/~vivienp/diss/>
- Schaer, Philipp: User interface design for search-term recommendation and interactive query expansion services. - 8th European Networked Knowledge Organization Systems (NKOS) Workshop 2009 URL: <http://www.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2009/programme.html>
- White, Howard D. (1981): 'Bradfordizing' search output: how it would help online users. In: Online Review 5, No. 1, pp. 47-54

